What do we really know about fiscal sustainability in the EU? A panel data diagnostic

Afonso, António; Rault, Christophe

Rev World Econ (2010) 145:731–755
DOI 10.1007/s10290-009-0034-1

ORIGINAL PAPER

What do we really know about fiscal sustainability in the EU? A panel data diagnostic

António Afonso · Christophe Rault

Published online: 6 November 2009

Abstract We assess the sustainability of public finances in the EU-15 over the period 1970–2006 using stationarity and cointegration analysis. Specifically, we use panel unit root tests of the first and second generation allowing in some cases for structural breaks. We also apply modern panel cointegration techniques developed by Pedroni (Oxf Bull Econ Stat 61(1):653–670, 1999; Econom Theory 20(3):597–625, 2004), generalized by Banerjee and Carrion-i-Silvestre (Cointegration in panel data with breaks and cross-section dependence, European Central Bank, Working Paper 591, 2006) and Westerlund and Edgerton (Econ Lett 97(3):185–190, 2007), to a structural long-run equation between general government expenditures and revenues. While estimations point to fiscal sustainability being an issue in some countries, fiscal policy was sustainable both for the EU-15 panel set, and within sub-periods (1970–1991 and 1992–2006).

Keywords Intertemporal budget constraint · Fiscal sustainability · EU · Panel unit root · Panel cointegration

A. Afonso
European Central Bank, Kaiserstraße 29, 60311 Frankfurt am Main, Germany
e-mail: antonio.afonso@europa.ecb.eu; aafonso@iseg.utl.pt

A. Afonso
Department of Economics, UECE—Research Unit on Complexity and Economics, ISEG/TULisbon—Technical University of Lisbon, R. Miguel Lupi 20, 1249-078 Lisbon, Portugal

C. Rault (✉)
LEO, University of Orléans, Rue de Blois, UMR 6221, B.P. 6739, 45067 Orléans Cedex 2, France
e-mail: christophe.rault@univ-orleans.fr; chrault@hotmail.com
URL: http://membres.lycos.fr/chrault/index.html

C. Rault
BEM, Bordeaux Management School, Bordeaux, France

JEL Classification C23 · E62 · H62 · H63

1 Introduction

The sustainability of public finances is a key policy issue for the European Union (EU). Within the EU fiscal framework, fiscal discipline is an important support for the implementation of monetary policy, particularly in the case of the EMU member countries. In EMU, the existence of sound fiscal policies is seen as a necessary objective for individual countries to pursue. It is not possible to exclude adverse responses from the financial markets when fiscal behaviour is deemed to be unsustainable. Indeed, the accumulation of government debt, following continued budgetary imbalances, may in the end trigger the need for higher long-term interest rates in order to place additional sovereign debt in the markets. Moreover, the Treaties governing the EU also require sustainable public finances. Countries are urged to comply with the budgetary requirements of EMU, by avoiding excessive deficits, keeping debt levels below the 60% of GDP reference value, and respecting the requirements of the stability and growth pact (SGP).

The aim of this paper is to examine the sustainability of public finances for the EU-15 countries (covering the EU Member States before the 1 May 2004 enlargement) by applying recent advances in the econometrics of non-stationary panel data methods.[1] The econometric literature on unit roots and cointegration testing has been expanding rapidly, and now distinguishes between the first generation tests developed on the assumption of cross-section independence (except for common time effects), and the second generation tests that allow, in a variety of forms and degrees, the dependence that might prevail across the different units in the panel. This question is crucial and responds to the complex nature of the interactions and dependencies that generally exist over time and across the individual units in the panel. For instance, observations on firms, industries, regions and countries tend to be cross-correlated as well as serially dependent. As pointed out by Breitung and Pesaran (2005), the problem of cross-section dependence is particularly difficult to deal with since it could arise for a variety of reasons, including spatial spillover effects, common unobserved shocks, social interactions, or a combination of these factors. In the context of our paper, cross-dependence can mirror possible changes in the behaviour of fiscal authorities related to the signing of the EU Treaty in Maastricht on 7 February 1992, with the setting up of the convergence criteria that urged the EU countries to consolidate public finances in the run-up to the EMU on 1 January 1999, when most EU legacy currencies were replaced by the euro, and in the context of the SGP since then.

Generally, fiscal sustainability is considered on a country basis and can usually only be restored by changing national fiscal policies. From a monetary policy point of view, fiscal policy in the current institutional setting of EMU must be considered a largely national competence and responsibility. Although, even if there is no

[1] The countries are: Belgium, Denmark, Germany, Ireland, Greece, Spain, France, Italy, Luxembourg, the Netherlands, Austria, Portugal, Finland, Sweden and the UK.

single fiscal policy in the EU, a panel sustainability analysis of public finances has to be seen as relevant in a context of EU countries seeking to pursue common and sound fiscal policy behaviour within the SGP framework. Possible cross-country dependence can be envisaged either in the run-up to EMU or, for example, via integrated financial markets. Indeed, with cross-country spillovers in government bond markets especially after the completion of the single EU-15 capital market from 1994 were to be expected, interest rates comovements inside the EU became also more noticeable.

To the best of our knowledge, few comparable studies have taken into account the possible cross-sectional dependence among countries when investigating the sustainability of public finances for the EU-15 countries. A few studies provide panel unit root and panel cointegration analysis in this context, notably Prohl and Schneider (2006), for eight OECD countries and Claeys (2007) for the EU (not allowing for cross-section dependence). Indeed, although the main analytical techniques used to analyse the sustainability of public finances have been stationarity tests for the stock of public debt and cointegration tests between government expenditures and government revenues, this has been mostly performed for individual countries, which sometimes poses the problem of relatively short time series.[2] This paper takes these results in the literature regarding the sustainability of public finances, and assesses them to see whether they still hold when more powerful cointegration techniques are employed in a panel framework.

Our econometric methodology uses two approaches for unit root testing: panel data integration tests of “first generation” (Im et al. 2003; Levin et al. 2002), which assume cross-sectional independence among panel units (except for common time effects); and panel data unit root tests of the “second generation” (Choi 2006; Moon and Perron 2004), which allow for more general forms of cross sectional dependency (not only limited to common time effects). We also implement panel cointegration techniques developed by Pedroni (1999, 2004), and generalised by Banerjee and Carrion-i-Silvestre (2006) and Westerlund and Edgerton (2007), to a structural long-run equation between general government expenditures and revenues. The advantages of panel data methods within the macro-panel setting include the use of data for which the spans of individual time series data are insufficient for the study of many hypotheses of interest. Other benefits include better properties of the testing procedures when compared to more standard time series methods, and the fact that many of the issues studied, such as convergence, purchasing power parity or the sustainability of public finances, naturally lend themselves to being studied in a panel context.

The remainder of the paper is organised as follows. In Sect. 2 we briefly review the analytical framework of public finance sustainability. In Sect. 3 we present a brief overview of our fiscal data. In Sect. 4 we perform the stationarity analysis of the fiscal series. In Sect. 5 we report the cointegration results for the general government expenditure and revenue series. Finally, Sect. 6 concludes the paper.

[2] Examples of empirical tests of fiscal sustainability on an individual country basis are provided, for instance, by Hamilton and Flavin (1986), Trehan and Walsh (1991), Wilcox (1989), Hakkio and Rush (1991), Tanner and Liu (1994), Quintos (1995), Haug (1991), Ahmed and Rogers (1995), Payne (1997), Bohn (1998), Fève and Hénin (2000), Uctum and Wickens (2000), Bergman (2001) and Afonso (2005).

2 The analytical framework of public finance sustainability

In the beginning of the 1920s, when writing about the public debt problem faced by France, Keynes (1923) highlighted the need for the French government to conduct a sustainable fiscal policy in order to satisfy its budget constraint. Keynes stated that the absence of sustainability would be evident when "the State's contractual liabilities (…) have reached an excessive proportion of the national income" (p. 54).

In modern terms, the sustainability of public finances is challenged when the government debt-to-GDP ratio reaches an excessive value. There is a problem of sustainability when the government revenues are not enough to keep on financing the costs associated with the new issuance of public debt or, again in Keynes words, when "it has become clear that the claims of the bond-holders are more than the tax payers can support" (p. 55). At that point the government will have to take measures that restore the sustainability of fiscal policy, meaning that the State "must come in due course to some compromise between increasing taxation, and diminishing expenditure, and reducing what (…) [it] owe[s]" (p. 59).

From an analytical perspective, the issue of fiscal policy sustainability can be presented in a straightforward way with the so-called present value borrowing constraint (PVBC). In order to derive the PVBC of a single country, the flow government budget constraint for a given period t can be written as

$$G_t + (1 + r_t)B_{t-1} = R_t + Bt, \tag{1}$$

where G is the primary government expenditure, R is the government revenue, B is the government debt, and r is the real interest rate.[3] Rewriting (1) for the subsequent periods, and recursively solving that equation leads to the following intertemporal budget constraint:

$$B_t = \sum_{s=1}^{\infty} \frac{R_{t+s} - G_{t+s}}{\prod_{j=1}^{s}(1 + r_{t+j})} + \lim_{s\to\infty} \prod_{j=1}^{s} \frac{B_{t+s}}{(1 + r_{t+j})}. \tag{2}$$

When the second term from the right-hand side of Eq. (2) is zero, the present value of the existing stock of public debt will be identical to the present value of future primary surpluses. For empirical purposes it is useful to make several algebraic modifications to Eq. (1). Assuming that the real interest rate is stationary, with mean r, and defining

$$E_t = G_t + (r_t - r)B_{t-1}, \tag{3}$$

it is possible to obtain the following PVBC:

$$B_{t-1} = \sum_{s=1}^{\infty} \frac{1}{(1 + r)^{s+1}}(R_{t+l} - E_{t+s}) + \lim_{s\to\infty} \frac{B_{t+s}}{(1 + r)^{s+1}}. \tag{4}$$

A sustainable fiscal policy needs to ensure that the present value of the stock of public debt, the second term of the right-hand side of (4), goes to zero in infinity,

[3] For the validation of theoretical results, the real interest rate is sometimes assumed in the literature to be stationary, but this is a much more difficult assumption for the nominal interest rate.

constraining the debt to grow no faster than the real interest rate. In other words, it implies imposing the absence of Ponzi games and the fulfilment of the intertemporal budget constraint. Faced with this transversality condition, the government will have to achieve future primary surpluses whose present value adds up to the current value of the stock of public debt.[4]

It is also worth noting that the hypothesis of fiscal policy sustainability is related to the condition that the trajectory of the main macroeconomic variables is not affected by the choice between the issuance of public debt and the increase in taxation. Under such conditions, it would therefore be irrelevant how the deficits are financed, which also implies the assumption of the Ricardian Equivalence hypothesis.[5]

In addition, one can also derive the solvency condition, with all the variables defined as a percentage of GDP.[6] The PVBC, with the variables expressed as ratios of GDP, with y being the real GDP growth rate, and neglecting for presentation purposes seigniorage revenues, is then written as

$$\frac{B_t}{Y_t} = \frac{(1+r_t)B_{t-1}}{(1+y_t)Y_{t-1}} + \frac{G_1}{Y_1} - \frac{R_1}{Y_1}. \tag{5}$$

Assuming the real interest rate to be stationary, with mean r, and considering also constant real GDP growth, the budget constraint is then given by

$$b_{t-1} = \sum_{s=0}^{\infty} \left(\frac{1+y}{1+r}\right)^{(s+1)} [\rho_{t+s} - e_{t+s}] + \lim_{s\to\infty} b_{t+s} \left(\frac{1+y}{1+r}\right)^{(s+1)} \tag{6}$$

with $b_t = B_t/Y_t$, $e_t = E_t/Y_t$ and $\rho_t = R_t/Y_t$. When $r > y$, it is necessary to introduce a solvency condition, given by $\lim_{s\to\infty} b_{t+s}\left(\frac{1+y}{1+r}\right)^{(s+1)} = 0$, in order to bound public debt growth.[7] This yields the familiar result that fiscal policy will be sustainable if the present value of the future stream of primary surpluses, as a percentage of GDP, matches the "inherited" stock of government debt. In a similar fashion, looking at the US after the end of the Second World War, Domar (1944) pointed out that it would be possible to sustain successive primary budget deficits as long as the real growth rate surpasses the real interest rate ($y > r$).

A common practice in the literature is to investigate past fiscal data to see if government debt follows a stationary process or to establish if there is cointegration between government revenues and government expenditures.[8] Recalling the PVBC in Eq. (4), it is possible to ascertain empirically the absence of Ponzi games by testing the stationarity of the first difference of the stock of public debt, using unit

[4] McCallum (1984) discusses whether this is a necessary condition to obtain an optimal growth trajectory for the stock of public debt.

[5] Afonso (2008) provides evidence of overall Ricardian behaviour on the part of EU-15 governments.

[6] For instance, Hakkio and Rush (1991) suggest that an analysis based on ratios (to GDP) is more appropriate for growing economies.

[7] This implies that the growth rate of the debt-to-GDP ratio should be less than the factor $((1 + y)/(1 + r))^{(s+1)}$.

[8] Hamilton and Flavin (1986) first used these procedures. See also Trehan and Walsh (1991) and Hakkio and Rush (1991).

root tests both at the country level and for a European panel. It is also possible to assess fiscal policy sustainability through cointegration tests. The implicit hypothesis concerning the real interest rate, with mean r, is also stationarity. Using again the auxiliary variable $E_t = G_t + (r_t - r)B_{t-1}$, and the additional definition $GG_t = G_t + r_t B_{t-1}$, the intertemporal budget constraint may also be written as

$$GG_t - R_t = \sum_{s=0}^{\infty} \frac{1}{(1+r)^{s-1}}(\Delta R_{t+s} - \Delta E_{t+s}) + \lim_{s \to \infty} \frac{B_{t+s}}{(1+r)^{s+1}} \qquad (7)$$

and with the no-Ponzi game condition, GG_t and R_t must be cointegrated variables of order one for their first differences to be stationary.

Assuming that R and E are non-stationary variables, and that the first differences are stationary variables, this implies that the series R and E in levels are I(1). Then, for Eq. (7) to hold, its left-hand side will also have to be stationary. If it is possible to conclude that GG and R are integrated of order 1, these two variables should be cointegrated with cointegration vector $(1, -1)$ for the left-hand side of Eq. (7) to be stationary.

The procedure to assess the sustainability of the intertemporal government budget constraint therefore involves testing the following cointegration regression: $R_t = a + bGG_t + u_t$. If the null of no cointegration, i.e. the hypothesis that the two I(1) variables are not cointegrated, is rejected (with a high-test statistic), this implies that one should accept the alternative hypothesis of cointegration. For that result to hold true, the series of the residual u_t must be stationary, and should not display a unit root.

Hakkio and Rush (1991) also demonstrate that if GG and R are non-stationary variables in levels, the condition $0 < b < 1$ is a sufficient condition for the budget constraint to be obeyed. However, when government revenues and expenditures are expressed as a percentage of GDP (or in per capita terms), it is necessary to have $b = 1$ in order for the trajectory of the government debt to GDP ratio not to diverge in an infinite horizon.[9] In terms of our subsequent empirical analysis, we will assess the stationarity of government debt, a sufficient but not necessary condition for fiscal sustainability, and the existence of cointegration between government revenues and expenditures, a necessary condition for fiscal sustainability.

3 Fiscal data overview

All data are taken from the European Commission AMECO (Annual Macro-Economic Data) database, covering the period 1970–2006 for the EU-15 countries.[10] Table 1 reports summary statistics for our main fiscal variables.

[9] Quintos (1995), Ahmed and Rogers (1995) and Bergman (2001) discuss the necessary conditions for sustainability in terms of the order of integration of public debt.

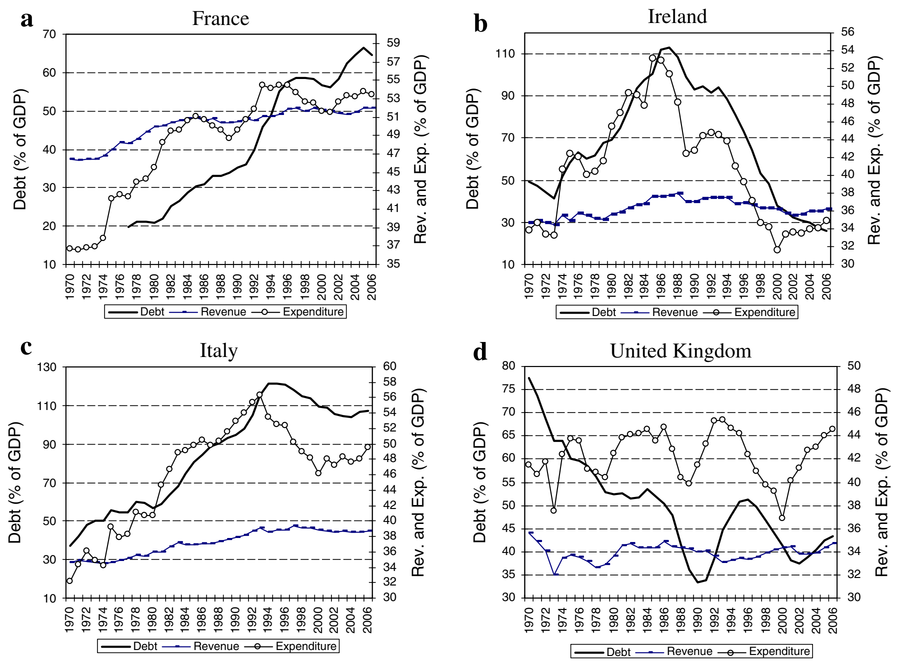
[10] AMECO codes: GDP at current market prices, .1.0.0.0.UVGD; gross domestic product, at 2000 market prices, .1.1.0.0.OVGD; general government consolidated gross debt, excessive deficit procedure (based on ESA 1995) and former definition (linked series) (% of GDP); .1.0.319.0.UDGGL, .1.0.319.0.UDGGF; general government debt (level), .1.0.0.0.UDGGL, .1.0.0.0.UDGGF; general government total expenditure (% of GDP), .1.0.319.0.UUTGE, .1.0.319.0.UUTGF; general government total revenue (% of GDP), .1.0.319.0.URTG, .1.0.319.0.URTGF; general government interest payments (% of GDP), .1.0.319.0.UYIG, .1.0.319.0.UYIGF (database updated on 04/05/2007).

Table 1 Statistical summary for fiscal variables (% of GDP, 1970–2006)

Country	Government debt				Primary balance			
	Mean	Max.	Min.	*n*	Mean	Max.	Min.	*n*
Austria	48.0	67.9	16.7	37	0.9	3.5	−2.0	37
Belgium	97.9	133.4	54.3	37	2.0	6.8	−4.8	37
Denmark	48.3	80.1	6.2	36	4.5	11.6	−3.0	37
Finland	26.6	57.8	6.1	36	4.0	9.7	−3.3	37
France	42.3	66.6	19.8	30	0.2	1.9	−2.3	37
Germany	42.5	67.9	18.0	37	0.2	2.8	−4.1	37
Greece	67.2	114.0	17.5	30	−0.7	5.0	−6.7	37
Ireland	67.5	112.9	25.8	37	0.8	6.6	−7.3	37
Italy	84.9	121.5	37.4	37	−0.7	6.6	−6.7	37
Luxembourg	9.3	20.3	4.1	37	2.6	6.4	−1.6	37
Netherlands	60.6	78.5	39.6	32	1.7	5.0	−1.3	37
Portugal	47.7	67.4	14.2	34	−0.4	3.9	−7.4	37
Spain	37.3	66.8	11.8	32	0.0	3.1	−4.4	37
Sweden	49.2	73.2	24.6	34	4.0	10.3	−5.6	37
United Kingdom	49.9	77.4	33.4	37	1.1	6.8	−4.8	37
Country	**Government revenue**				**Government expenditure**			
	Mean	Max.	Min.	*n*	Mean	Max.	Min.	*n*
Austria	48.0	52.5	38.3	37	50.1	56.7	37.1	37
Belgium	46.3	51.1	38.1	37	51.6	62.1	40.2	37
Denmark	52.9	58.1	44.0	37	52.6	60.6	39.5	37
Finland	48.9	57.1	33.6	37	46.5	64.7	29.5	37
France	46.2	50.9	37.1	37	48.4	54.5	36.5	37
Germany	44.3	46.6	39.6	37	46.6	49.9	39.1	37
Greece	34.0	47.0	22.5	37	40.3	52.0	22.6	37
Ireland	36.5	43.6	29.2	37	40.9	53.2	31.6	37
Italy	38.7	47.6	27.9	37	46.2	56.3	32.1	37
Luxembourg	40.4	44.4	27.8	35	38.5	45.2	25.3	35
Netherlands	48.5	53.8	41.2	37	51.0	59.2	42.7	37
Portugal	32.6	43.5	20.6	37	36.9	47.8	18.6	37
Spain	32.8	40.1	20.9	37	35.2	46.6	20.3	37
Sweden	57.4	62.3	46.0	37	57.6	72.4	41.8	37
United Kingdom	39.8	44.1	34.9	37	42.3	45.4	36.9	37

Source European Commission AMECO database

In the period 1970–2006 the highest government debt-to-GDP ratios were recorded in Belgium, Italy, Greece and Ireland, related to high budget deficits incurred by those countries, and resulted notably in the pushing up of interest payments. The government expenditure-to-GDP ratios ranged overall between some 20 and 70%, with the lower values being recorded in the beginning of the period,



Source: European Commission AMECO database.

Fig. 1 Fiscal variables for selected countries. *Source* European Commission AMECO database

while the government revenue-to-GDP ratios were in the interval between 20 and 60%. Additionally, visual inspection of the revenue and expenditure time series as a ratio of GDP, as exemplified in Fig. 1 for selected countries, and in advance of the subsequent econometric analysis, may help to assess sustainability issues in individual cases.

4 Stationarity analysis of fiscal series

In this section we study the stationarity of the fiscal series in our country panel, specifically the stock of government debt in real terms and the ratios to GDP of government revenue and government expenditure, using several panel unit root tests, which allow notably for cross-country independence and dependence.[11]

4.1 First generation panel unit root tests (cross-country independence)

In this sub-section, we implement the following "first generation" panel data unit root tests (Im et al. 2003; Levin et al. 2002). First, we used the test proposed by Im

[11] Note that to make the analysis robust, we also compared the results of panel data unit root tests with those obtained with individual unit root tests (see Sect. 7 for a summary of the results). For complete details on this comparison see the working paper version in Afonso and Rault (2007).

Table 2 Summary of panel data unit root tests for the first difference of the stock of government debt, constant prices (1970–2006)

Method	Statistic	*P*-value[a]	Cross-sections	Obs.
Null: unit root (assumes common unit root process)				
Levin, Lin and Chu *t*-stat.	−1.92991	0.0268	15	494
Null: unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat.	−3.18952	0.0007	15	494

Automatic selection of lags based on SIC. Newey–West bandwidth selection using a Bartlett kernel

[a] The tests assume asymptotic normality

et al. (2003, hereafter IPS), which has been widely implemented in empirical research due to its rather simple methodology and alternative hypothesis of heterogeneity. This test assumes cross-sectional independence among panel units (except for common time effects), but allows for heterogeneity in the form of individual deterministic effects (constant and/or linear time trend), and heterogeneous serial correlation structure of the error terms. Tables 2, 3 and 4 report the results of the IPS test for the government debt, and for the revenue and expenditure ratio series. In order to facilitate comparisons, we also provide the results of the panel unit root tests of Levin et al. (2002).

Concerning the first difference of the stock of government debt, the results given by the panel data unit root tests are more concomitant than those provided by the standard (individual) unit root ones. Indeed, at the 5% level of significance, the two panel data tests reveal that the null unit root hypothesis can be rejected at the 5% level for EU-15 countries (see Table 2), thus supporting the stationarity of the change in the stock of government debt and hence the non-rejection of the solvency condition for the overall country sample.[12]

As far as the general government revue-to-GDP ratio is concerned, the two panel data tests produce significant evidence in favour of their integration of order one for all EU-15 countries at the 5% level of significance (see Table 3). In other words, the non-stationarity of the revenue-to-GDP ratio cannot be rejected. Finally, and according to Table 4, the general government expenditure-to-GDP ratio also appears to have a unit root for all countries at the 5% level of significance if one refers to the results of the two panel data unit root tests.

However, as shown by several authors (notably O'Connell 1998; and Banerjee et al. 2004, 2005), the assumption of cross-sectional dependence limited to the case of common time effects on which the asymptotic results of the IPS's procedure relies (like most panel data unit root tests of "the first generation", including Levin et al. 2002) is often unrealistic and can be at odds with economic theory and

[12] A common feature of the panel tests mentioned above is that they maintained the null hypothesis of a unit root in all panel members. Therefore, their rejection decision actually indicates that at least one panel member is stationary, with no information about how many series or which ones are stationary. This possibility for a mixed panel implies that some of the members may be stationary while others may be non-stationary (see Taylor and Sarno 1998 and Taylor and Taylor 2004 for further details).

Table 3 Summary of panel data unit root tests for general government revenue-to-GDP ratios (1970–2006)

Method	Statistic	P-value[a]	Cross-sections	Obs.
Null: unit root (assumes common unit root process)				
Levin, Lin and Chu t-stat.	−0.77258	0.2199	15	534
Null: unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat.	2.09943	0.9821	15	534

Automatic selection of lags based on SIC. Newey–West bandwidth selection using a Bartlett kernel

[a] The tests assume asymptotic normality

Table 4 Summary of panel data unit root tests for general government expenditure-to-GDP ratios (1970–2006)

Method	Statistic	P-value[a]	Cross-sections	Obs.
Null: unit root (assumes common unit root process)				
Levin, Lin and Chu t-stat.	−0.88260	0.1887	15	450
Null: unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat.	2.61169	0.9955	15	450

Automatic selection of lags based on SIC. Newey–West bandwidth selection using a Bartlett kernel

[a] The tests assume asymptotic normality

empirical results. Besides, as shown in two simulation studies by Banerjee et al. (2004, 2005), if panel members are cross-correlated or even cross-sectionally cointegrated, all these tests experience strong size distortions and limited power. This point is analytically confirmed by Lyhagen (2000) and Pedroni and Urbain (2001).

4.2 Second generation panel unit root tests (cross-country dependence)

As Breitung and Pesaran (2005) note, time series are contemporaneously correlated in many macroeconomic applications using country or regional data. Prominent examples of this are the analysis of purchasing power parity and output convergence (see for instance Pesaran 2004). However, the literature on how to model cross-sectional dependence in large panels is still developing. Cross-sectional dependence can arise due to a variety of factors, such as omitted observed common factors, spatial spillover effects, for example via integrated financial markets, unobserved common factors, or general residual interdependence, all of which could remain even when all observed and unobserved common effects have been taken into account. In the EU context, some possible cross-country dependence can be envisaged in the presence of a similar policy measures (i.e. in the run-up to EMU), coupled with similar fiscal behaviour (e.g. pursuing fiscal consolidation in the run-up to EMU and within the SGP framework), and cross-country spillovers in

Table 5 Results of Choi (2006) test (1970–2006)

	P_m-statistic	Z-statistic	L^*-statistic
First difference of the stock of public debt	0.000	0.000	0.000
General government revenue-to-GDP ratios	0.463	0.354	0.354
General government expenditure-to-GDP ratios	0.364	0.382	0.373

Note that the P_m-test is a modification of Fisher's (1932) inverse Chi-square tests, and rejects the null unit root hypothesis for positive large value of the statistics, and that the L^* is a logit test. The tests (Z and L^*) reject the null for large negative values of the statistics. The P-, Z- and L^*-tests converge under the null to a standard normal distribution as (N, $T \to \infty$) (see Choi 2006 for further details)

Note All figures reported in the table are p-values

government bond markets especially after the completion of the single EU-15 capital market from 1994 (stage 2 of EMU).[13]

For this reason, various recent studies have proposed panel unit root tests allowing for more general forms of cross-sectional dependency, e.g. Choi (2006), Moon and Perron (2004), and Phillips and Sul (2003). We have decided to investigate the presence of a unit root using two-second generation tests, namely Choi (2006) and Moon and Perron (2004), to whom we refer the reader for further details.[14] This last test in particular seems to show good size and power for different values of T and N and model specifications, according to the Monte Carlo experiments conducted by Gutierrez (2006).[15]

The results reported in Tables 5 and 6 indicate that the null unit root hypothesis cannot be rejected by the two tests at the 5% level for the government expenditure and revenue ratios, but can be rejected for the government debt for all EU-15 countries, which supports the initial results produced by the first generation panel data unit root tests. Furthermore, tests on the series in first differences confirm the hypothesis of stationarity for government expenditure and revenue ratios. Therefore, we may conclude that the general government revenue and expenditure-to-GDP ratios expressed in level are integrated of order 1 for all EU-15 countries, independently of the panel unit root tests considered, thereby demonstrating that the non-stationarity property of our revenue and expenditure series is a robust result.

[13] It should be noted that before carrying out the second generation panel unit-root tests that account for cross-section dependence, we have first implemented the simple test of Pesaran (2004) and have computed the CD statistic to test for the presence of such cross-section dependence in the data. This test is based on the average of pair-wise correlation coefficients of the OLS residuals obtained from standard augmented Dickey–Fuller regressions for each individual. Its null hypothesis is cross-sectional independence and is asymptotically distributed as a two-tailed standard normal distribution. The null hypothesis is always rejected regardless of the number of lags included in the augmented DF auxiliary regression (up to five lags) at the 5% level of significance. This confirms that the members of our panel are cross-sectionally correlated.

[14] Note that another possibility would be to use a procedure as the one advocated by Breuer et al. (2002) whereby unit root testing is conducted within a seemingly unrelated regression (SUR) framework. An advantage of this procedure is that the SUR framework is another useful way of addressing cross-sectional dependency.

[15] We are grateful to C. Hurlin for making available his Matlab codes to us.

Table 6 Results of Moon and Perron (2004) test (1970–2006)

	$t \times a$	$t \times b$
First difference of the stock of public debt	0.000	0.000
General government revenue-to-GDP ratios	0.526	0.541
General government expenditure-to-GDP ratios	0.382	0.434

The null hypothesis of the two tests proposed by Moon and Perron (2004) is the unit root for all panel units. Under the null H_0, they show that for (N, $T \to \infty$) with $N/T \to 0$, the statistics $t \times a$ and $t \times b$ have a standard normal distribution

Note All figures reported in the table are p-values

4.3 Panel unit root tests allowing for structural breaks

The presence of structural breaks in panel series data can induce behaviour similar to that of an integrated process, making it difficult to differentiate between a unit root and a stationary process with a regime shift. For this reason, the panel unit root tests in the previous section, such as the IPS test, may potentially suffer from a significant loss of power if structural breaks are present in the data.

In this section, we employ the panel data unit root test based on the Lagrangian multiplier (LM) principle developed by Im and Lee (2001), which is very flexible since it can be applied not only when a structural break occurs at a different time period in each time series, but also when the structural break occurs in only some of the time series. The proposed test is not only robust to the presence of structural breaks, but is also more powerful than the popular IPS test in the basic scenario where no structural breaks are involved. Furthermore, as reported by Im and Lee (2001), since the LM test loses little power by controlling for spurious structural breaks when they do not exist, this represents a reasonable strategy to control for breaks even when they are only at a suspicious level. Moreover, this panel LM test does not require the simulation of new critical values that depend on the number and location of breaks.[16]

In order to provide a robust analysis, we compare both univariate and panel LM unit root test results with and without a structural break. We begin with the Schmidt and Phillips (1992) univariate LM unit root test without any structural change. Then, we move to extensions that allow for one break, since our time series covers periods during which structural change may have occurred due to structural and institutional changes in the EU-15 countries. In addition to the Schmidt and Phillips (1992) no-break test, we employ the univariate test and the Lee and Strazicich (2003) minimum LM unit root tests with one break to determine the structural break point in each country. After determining the optimal break point, we employ the panel LM unit root test of Im and Lee (2001). For comparison, we also show the panel LM test results with no breaks.

To determine the optimal break point in the panel LM test, we utilize the univariate minimum LM unit root tests of Lee and Strazicich (2003). These tests are

[16] It should be noted that these tests assume cross-sectional independence among panel units.

Table 7 Panel LM unit root tests allowing for structural break for the first difference of the stock of government debt (1970–2006)

Country	Individual LM statistic without a break[a]	Lags	Individual LM statistic with a break[b]	Lags	Optimal break point
Austria	−4.420*	7	−4.707*	7	2003
Belgium	−2.246	8	−1.632	1	1995
Denmark	−2.288	2	−4.126*	7	2000
Finland	−1.945	8	−2.456	3	1993
France	−2.997	3	−3.718*	4	1993
Germany	−2.877	8	−3.075	8	1993
Greece	−3.213*	8	−2.099	8	2002
Ireland	−1.444	2	−2.683	5	1995
Italy	−4.404*	7	−4.905*	7	2003
Luxembourg	−1.449	5	−1.731	4	1997
Netherlands	−0.487	3	−0.868	3	1992
Portugal	−1.874	3	−2.132	3	2002
Spain	−1.599	1	−1.076	1	1993
Sweden	−2.129	1	−3.155	1	2000
United Kingdom	−2.142	4	−2.169	4	2002
Panel LM stat.[c]	−3.126*		−5.077*		

Notes As all tests are one-sided, a calculated statistic smaller than the critical value leads to the rejection of the null of a unit root. At 5% the critical value for the LM test without break is −3.06. At 5% the critical value for the minimum LM test with one break is −3.566

The critical value for the panel LM test (with or without breaks) is −1.645 with an asymptotic standard normal distribution

* denotes significance at the 5% level

[a] Schmidt and Phillips (1992) test; [b] Lee and Strazicich (2003) test; [c] Im and Lee (2001) test

comparable to the corresponding Dickey and Fuller-type endogenous break tests of Zivot and Andrews (1992). The performance of the LM test is comparable to or superior to these counterpart tests in terms of size and power. In addition, the LM unit root tests are not subject to spurious rejections under the null. In each test, the break point is determined endogenously from the data via a grid search by selecting the break where the value of the unit root test statistic is at its minimum. Using the minimum LM tests of Lee and Strazicich (2003), the unit root test statistic is estimated at each break point. The procedure is repeated over the time interval [0.1 T, 0.9 T] in order to eliminate end points, until the break is determined where the unit root t-test statistic is minimized. The optimal number of lags in each country is determined by sequentially examining the t-statistic for the last lag coefficient to see if it is significant at the approximate 5% level in an asymptotic normal distribution. We begin with the one-break LM test. If less than one break is significant, we employ the no-break LM unit root test. The corresponding LM unit root test statistic is then chosen after determining the optimal break point. After determining the

Table 8 Panel LM unit root tests allowing for structural break for general government revenue-to-GDP ratios (1970–2006)

Country	Individual LM statistic without a break[a]	Lags	Individual LM statistic with a break[b]	Lags	Optimal break point
Austria	−2.667	2	−2.957	2	1989
Belgium	−1.627	3	−2.313	3	1990
Denmark	−2.128	6	−2.467	6	1989
Finland	−3.901*	8	−3.806*	8	2001
France	−3.063	4	−4.205*	6	1995
Germany	−1.593	7	−2.492	8	1998
Greece	−1.292	0	−1.443	0	1992
Ireland	−0.916	5	−0.346	8	1997
Italy	−2.284	8	−3.950*	8	1991
Luxembourg	0.502	8	0.362	8	1992
Netherlands	−2.070	2	−4.168*	8	1992
Portugal	−1.674	0	0.105	8	1988
Spain	−1.928	0	−1.577	8	1983
Sweden	−0.595	6	−0.811	8	1991
United Kingdom	−3.156*	6	−2.141	5	1987
Panel LM stat.[c]	−0.292		−1.62		

Notes As all tests are one-sided, a calculated statistic smaller than the critical value leads to the rejection of the null of a unit root. At 5% the critical value for the LM test without a break is −3.06. At 5% the critical value for the minimum LM test with one break is −3.566

The critical value for the panel LM test (with or without breaks) is −1.645 with an asymptotic standard normal distribution

* denotes significance at the 5% level

[a] Schmidt and Phillips (1992) test; [b] Lee and Strazicich (2003) test; [c] Im and Lee (2001) test

appropriate unit root test statistic for each country, the panel LM test statistic is then calculated.[17]

The results are reported in Tables 7, 8 and 9, which respectively show the first difference of the stock of government debt at 2000 constant prices, and general government expenditure and revenue taken as a percentage of GDP. For the univariate LM test with no break, the unit root null can be rejected at the 5% level of significance in three countries for government debt (Austria, Greece and Italy), in two countries for government expenditure (Finland and the UK), and in two countries for government revenue (Denmark and Sweden). After allowing for a structural break, the univariate minimum LM test rejects the unit root null in four countries for government debt (Austria, Denmark, France and Italy), in four countries for government expenditure (Finland, France, Italy and the Netherlands), and cannot reject it for government revenue at the 5% level.

[17] We are grateful to J. Lee for providing us with the GAUSS codes, which we have adapted for our analysis, and that are available upon request.

Table 9 Panel LM unit root tests allowing for structural break for general government expenditure-to-GDP ratios (1970–2006)

Country	Individual LM statistic without a break[a]	Lags	Individual LM statistic with a break[b]	Lags	Optimal break point
Austria	−1.627	3	−1.253	2	1981
Belgium	−2.128	6	−1.855	6	1991
Denmark	−3.901*	8	−2.055	8	1985
Finland	−3.063	4	−1.935	7	1981
France	−1.593	7	−1.712	2	1993
Germany	−1.292	0	−1.553	6	1993
Greece	−0.916	5	−2.779	7	1994
Ireland	−2.284	8	−1.487	7	1990
Italy	0.502	8	−2.372	7	2000
Luxembourg	−2.070	2	0.234	6	2003
Netherlands	−1.674	0	−1.394	7	1985
Portugal	−1.928	0	−1.966	8	2000
Spain	−0.595	6	−1.898	5	1986
Sweden	−3.156*	6	−1.203	1	1993
United Kingdom	−1.411	2	−1.326	7	2000
Panel LM stat.[c]	0.212		0.999		

Notes As all tests are one-sided, a calculated statistic smaller than the critical value leads to the rejection of the null of a unit root. At 5% the critical value for the LM test without a break is −3.06. At 5% the critical value for the minimum LM test with one break is −3.566

The critical value for the panel LM test (with or without breaks) is −1.645 with an asymptotic standard normal distribution

* denotes significance at the 5% level

[a] Schmidt and Phillips (1992) test; [b] Lee and Strazicich (2003) test; [c] Im and Lee (2001) test

Without allowing for structural breaks, the panel LM test statistic is −3.126 for the stock of real government debt series clearly indicating that the unit root null can be rejected at the 5% level of significance, due to increased power from panel data (see Table 7). In addition, after allowing for structural breaks, the panel test statistic of −5.95 strongly rejects the unit root null at the 5% level. These results clearly demonstrate the gain in power from combining structural breaks with panel data. Since the panel LM test statistic is calculated using the average test statistic of all countries, it is possible that the panel results are due to a small number of outliers having a relatively large impact.

Examination of the univariate test statistics (with breaks) for each country reveals that Austria, Denmark, France and Italy might qualify as such an outlier, as they are the only four countries that reject the unit root null at the 5% level. In order to see if our panel results are robust to a possible outlier effect, we therefore recalculated the panel LM test statistic (with breaks) omitting these four countries. The resulting panel test statistic of −3.62 continues to reject the unit root null at the 5% level of significance, thus firmly supporting our hypothesis that the panel test results are not due to outliers.

Table 10 Summary of stationarity tests, 5% level of significance (H_0: unit root, non-stationarity, for most cases)

Set of results	First difference of stock of real government debt (2000 constant prices)		
1	Panel unit root 1st generation tests, country independence	Levin et al. (2002), Im et al. (2003): no unit root	
2	Panel unit root 2nd generation tests, country dependence	Choi (2006): no unit root	Moon and Perron (2004): no unit root
3	Individual LM unit root tests	Schmidt and Phillips (1992), no breaks, no unit root: AT, GR, IT	Lee and Strazicich (2003), with breaks, no unit root: AT, DK, FR, IT
4	Panel LM unit root tests	Im and Lee (2001), no breaks: no unit root	Im and Lee (2001), with breaks: no unit root
		General government revenue (% of GDP)	General government expenditure (% of GDP)
5	Panel unit root 1st generation tests, country independence	Levin et al. (2002), Im et al. (2003): unit root, non-stationarity	Levin et al. (2002), Im et al. (2003): unit root, non-stationarity
6	Panel unit root 2nd generation tests, country dependence	Choi (2006) and Moon and Perron (2004): unit root	Choi (2006) and Moon and Perron (2004): unit root
7	Individual LM unit root tests	Schmidt and Phillips (1992), no breaks, no unit root: FI, UK Lee and Strazicich (2003), with breaks, no unit root: FI, FR, IT, NL.	Schmidt and Phillips (1992), no breaks, no unit root: DK, SW Lee and Strazicich (2003), with breaks, no unit root: reject for all countries
8	Panel LM unit root tests	Im and Lee (2001), no breaks: unit root	Im and Lee (2001), with breaks: unit root

AT Austria, *DE* Germany, *DK* Denmark, *FI* Finland, *FR* France, *GR* Greece, *IR* Ireland, *IT* Italy, *LU* Luxembourg, *NL* Netherlands, *PT* Portugal, *SW* Sweden, *UK* United Kingdom

Concerning the general government expenditure and the revenue series taken as a percentage of GDP, it appears that the panel LM test statistics with or without a break cannot reject the null unit root hypothesis at the 5% level of significance, thus providing strong evidence in favour of a unit root in these two EU-15 country series.

Overall, our findings using panel data unit root tests that allow for structural breaks support the previous results of first and second generation panel data unit root tests, leading us to conclude that the stock of government debt series is integrated of order zero (indicating that the solvency condition would be satisfied for the EU-15 countries), and that the general government expenditure and the revenue series are integrated of order one. These findings are summarized in Table 10.

5 Cointegration between government expenditure and revenue ratios

After having confirmed the non-stationarity of our series of government revenue and expenditure for the EU-15 as a whole, in particular if one refers to the panel data unit root tests of the previous section, it is natural to test the existence of a structural long-run relationship between both series. This is the procedure we use in this section to assess fiscal sustainability on the basis of the intertemporal budget constraint as given in (7).

Compared to panel unit root tests, the analysis of cointegration in panels is still at an early stage of development. So far, the focus of the panel cointegration literature has been on residual-based approaches, although there have been a number of attempts to develop system approaches as well. As is the case for panel unit root tests, panel cointegration tests are based on homogeneous and heterogeneous alternatives. The residual-based tests were developed to ward against the spurious regression problem that can arise in panels when dealing with I(1) variables. Such tests are appropriate when it is a priori known that at most there can be only one within-group cointegration in the panel. Notable contributions to this strand of the literature include Pedroni (1999, 2000, 2004), and more recently Westerlund and Edgerton (2007).

The computation of the Pedroni test statistics assumes cross-sectional independence across individual units (apart from common time effects), an assumption that, as we have already mentioned, is probably absent for many macroeconomic time series. To take into account the possible cross-sectional dependence when carrying out the cointegration analysis, we decided to compute the bootstrap distribution of Pedroni's test statistics, thereby generating data-specific critical values. As in Banerjee and Carrion-i-Silvestre (2006), we have of course not used the seven statistics proposed by Pedroni (1999, 2004) to test the null hypothesis of no cointegration using single equation methods based on the estimation of static regressions. These statistics can also be grouped into either parametric or non-parametric statistics, depending on the way that autocorrelation and endogeneity bias are accounted for. In our study, we are only concerned with the parametric version of the statistics, i.e. the normalized bias and the pseudo *t*-ratio statistics, and with the ADF test statistics in particular. These test statistics are defined by pooling the individual tests, so that they belong to the class of between-dimension test statistics (see Pedroni, 1999, 2004 for further details).

As Banerjee and Carrion-i-Silvestre (2006) stress, some caution is required concerning the method used to bootstrap cointegration relationships, since not all available procedures lead to consistent estimates. In this regard, we have followed Phillips (2001), Park (2002) and Chang et al. (2006) in using a modified version of the sieve bootstrap described in Banerjee and Carrion-i-Silvestre (2006).[18]

Table 11 reports the results of the panel data cointegration tests developed by Pedroni (1999, 2004) both using conventional (asymptotic) critical values (as per Pedroni, 1999) and bootstrap critical values. We present the results for the entire sample period, 1970–2006, and for two sub-periods, 1970–1991 and 1992–2006, in order to assess whether different fiscal realities and behaviour can be detected for more recent years in the EU, notably after the signing of the Maastricht Treaty with the setting up of the fiscal convergence criteria.

[18] We are grateful to A. Banerjee and J. Carrion-i-Silvestre for providing us with their GAUSS codes (for a detailed discussion of the method used, see the end of the paper).

Table 11 Panel cointegration test results between government revenue and expenditure (Pedroni 1999, 2004)

	ADF-stat.	*P*-value	Bootstrap distribution		
			1%	5%	10%
Period 1970–2006					
Model with no deterministic component	−4.38	0.00	−4.88	−4.01	−3.52
Model with a constant term	−3.19	0.00	−4.25	−3.31	−2.82
Model including a time trend	−4.04	0.00	−5.62	−4.70	−4.03
Period 1970–1991					
Model with no deterministic component	−5.93	0.00	−7.63	−6.31	−5.63
Model with a constant term	−7.38	0.00	−6.68	−5.40	−4.72
Model including a time trend	−3.50	0.00	−7.56	−6.69	−5.09
Period 1992–2006					
Model with no deterministic component	−2.93	0.00	−6.78	−5.53	−4.87
Model with a constant term	−1.79	0.03	−7.78	−6.32	−5.62
Model including a time trend	−5.79	0.00	−9.22	−7.76	−6.98

Notes The bootstrap is based on 2000 replications

As the tests are one-sided, a calculated statistic smaller than the critical value leads to the rejection of the null hypothesis of no cointegration

For the period 1970–2006, using conventional asymptotic critical values (−1.65 at 5%) calculated under the assumption of cross-sectional independence (reported in Pedroni 1999, and extracted from the standard normal distribution), the null hypothesis of no cointegration between government revenue and expenditure ratios is always rejected by the test statistics, irrespective of whether the model includes a constant or a linear trend. However, if we consider bootstrap critical values (which are valid if there is some dependence among individuals), the conclusions of the test are less straightforward, and instead crucially depend on the level of significance chosen. Indeed, at the 10% level of significance, the null hypothesis of no cointegration is still rejected by the data, but an opposite result is obtained at the 5% level of significance for a model including either a constant or a linear trend. Finally, retaining a 10% level of significance, we conclude that a long-run relationship exists between government revenue and expenditure for the set of EU-15 countries, whatever the specification of the deterministic component.

We then investigated the robustness of the previous results, implementing panel data cointegration tests for the two sub-periods 1970–1991 and 1992–2006. The results are easier to interpret and provide econometric elements that justify this split on the basis of economic and institutional grounds, as two different types of behaviour now emerge from the cointegration tests (see Table 11).

First, concerning the 1970–1991 period, if one considers a model with a constant term, a statistical cointegration relationship clearly exists between government revenue and expenditure ratios, irrespective of whether one considers the (asymptotic) *p*-value or bootstrap critical values at 1, 5 or 10%. The opposite result is however obtained for a model including a time trend, independently of the

Table 12 Panel cointegration test results between government revenue and expenditure (Westerlund and Edgerton 2007, the null hypothesis of the tests is cointegration between government revenue and expenditure)

	LM-stat.	Asymptotic *p*-value	Bootstrap *p*-value
Period 1970–2006			
Model with a constant term	7.08	0.00	0.02
Model including a time trend	3.90	0.00	0.02
Period 1970–1991			
Model with a constant term	0.63	0.26	0.44
Model including a time trend	2.10	0.01	0.02
Period 1992–2006			
Model with a constant term	1.37	0.08	0.16
Model including a time trend	3.22	0.00	0.19

Note the bootstrap is based on 2,000 replications

critical values used (asymptotic or bootstrap). Finally, intermediate results are obtained for a model with no deterministic component, for which a long-run statistical relationship between government revenue and expenditure ratios only exists with the 10% bootstrap critical value.

Second, the results do not seem to confirm the existence of a cointegration relationship for the period 1992–2006 between government revenue and expenditure ratios in the EU-15 panel data set. This result is valid for any specification of the deterministic component considered, and is robust to the critical value used (asymptotic or bootstrap) for the conventional levels of significance. In this context, we should recall that after the beginning of the new millennium, the EU faced an economic recession (mirroring the beginning of the 1990s), with several countries entering into an excessive deficit procedure (EDP) situation within the fiscal framework of the SGP. The reason why some countries faced an EDP depended, to some extent, on the difficulties encountered in implementing sound fiscal policies in “good times” and thus the lack of budgetary manoeuvre in the recession period. Such developments may explain the different results regarding fiscal sustainability obtained in our analysis for this more recent period.

In order to assess the robustness of our findings, we also implemented the bootstrap panel cointegration test proposed by Westerlund and Edgerton (2007). Unlike the panel data cointegration tests of Pedroni (1999, 2004), here the null hypothesis is now cointegration. This new test relies on the popular Lagrange multiplier test of McCoskey and Kao (1998), and permits correlation to be accommodated both within and between the individual cross-sectional units. In addition, the bootstrap suggested by Westerlund and Edgerton (2007) is based on the sieve-sampling scheme, and has the appealing advantage of significantly reducing the distortions of the asymptotic test.[19] The results reported in Table 12 for a model including either a constant term or a linear trend clearly indicate the absence of a cointegrating relationship between government revenue and

[19] We are grateful to J. Westerlund for making available his GAUSS codes to us.

expenditure since with an asymptotic p-value of 0.00, the null hypothesis of cointegration is always rejected. This result is only marginally modified if one refers to the bootstrap critical value, indicating that for a significant level higher than 2%, the null hypothesis is still rejected. Hence at the conventional 5 and 10% levels of significance, we can conclude that there is no cointegrating relationship between government revenue and expenditure for the EU-15 panel data set.

Interestingly, performing the panel data cointegration tests for the two sub-periods 1970–1991 and 1992–2006 produces strong evidence in favour of the existence of a cointegration relationship between government revenue and expenditure ratios for the model with a constant term, with bootstrap p-values of 44% for the period 1970–1991, and 16% for the period 1992–2006. Hence, the necessary condition for public finance sustainability, i.e. the existence of a cointegration relationship between government revenue and expenditure, seems to be verified for the two sub-periods using this bootstrap panel cointegration test.

We further investigated whether public finances were sustainable for the model including a constant term, following the panel fully modified OLS approach developed in Pedroni (1996, 2000) and using a t-statistic to test whether the panel cointegration coefficient of the general government expenditure-to-GDP ratios is equal to one or not in the cointegrating regression where the government revenue-to-ratio is the dependent variable. For the period 1970–2006, the calculated t-statistic of 5.03 is above the tabulated critical values extracted from the normal distribution (1.96 and 2.33, respectively at the 5 and 1% levels of significance). The confidence intervals for this coefficient, at the 5% level of significance, [1.023; 1.136], confirm that the value of the coefficient is likely to be higher than one. For the two sub-periods, the 5% confident intervals for the coefficient are respectively [0.868; 1.072] for the period 1970–1991, and [0.678; 0.841] for the period 1992–2006. This therefore indicates that the coefficient in the cointegration relation is likely to be equal to one for the period 1970–1991, which provides evidence of the sustainability of public finances in that period.

Finally, we also tested, along the lines of MacDonald (1992), the possibility of cointegration between the primary balance ratio and the government debt-to-GDP ratio, which represents a possible avenue for assessing the sustainability of public finances, provided that both series are I(1) processes. However, the panel unit root tests for those series, as reported in sect. 8, show that while the government debt-to-GDP ratio is indeed I(1), the primary balance ratio is I(0), which thus excludes the possibility of the existence of a cointegration relationship between these two series.[20]

6 Conclusion

This paper has drawn on recent advances in the econometrics of non-stationary panel data methods to assess the sustainability of public finances for the EU-15

[20] Similar results, not reported here, are obtained with the implementation of the panel data tests of the second generation by Moon and Perron (2004) and Choi (2006).

countries in the period 1970–2006. Starting from the present value borrowing constraint of governments, we investigate past fiscal data to see if the stock of real government debt follows a stationary process, or if there is cointegration between government revenue and government expenditure as a percentage of GDP.

The econometric methods used in the paper to assess the sustainability of public finances in the EU-15 rest upon (1) first generation panel data integration tests that assume cross-sectional independence among panel units (apart from common time effects); (2) two-second generation panel data unit root tests that relax the assumption of cross-sectional independence; (3) panel data unit root tests that enable to accommodate structural breaks, and (4) the panel data cointegration tests developed by Pedroni (1999, 2004) and generalized by Banerjee and Carrion-i-Silvestre (2006), and the bootstrap panel cointegration test by Westerlund and Edgerton (2007).

The results from these panel unit root tests, allowing for structural breaks, support the results of both the first and second generation panel data unit root tests, leading us to conclude that the first difference of the stock of real government debt series is integrated of order zero, thus indicating that the solvency condition would be satisfied for EU-15 countries, which is a necessary condition for fiscal policy sustainability. Moreover, our results also show that general government expenditure and revenue ratios are integrated of order one.

Even if the results of the analysis may question fiscal sustainability in some cases when taken individually, it is nevertheless true that the tests point to the solvency of government public finances when considering the EU-15 panel data set. Naturally, this is an obvious advantage of the panel approach, since the time series dimension of the data is not that long for individual countries. Even if there is no single fiscal policy in the EU, the panel sustainability of public finances indicated by our results is relevant in a context of EU countries seeking to pursue sound fiscal policy behaviour within the SGP framework. Nevertheless, what we can also conclude from our analysis is that for some particular cases sustainability will not be attained if past fiscal behaviour is to be kept unchanged in the future. For instance, and as we saw, the solvency condition, on the basis of the stationarity tests of government debt, was satisfied for roughly half of the 15 EU countries: Austria, Finland, France, Germany, the Netherlands, the UK and Sweden. This set of countries is even smaller once we take into account the existence of structural breaks in the series.

Interestingly, the panel cointegration results for the entire 1970–2006 period allow us to draw the conclusion that a long-run relationship does exist between general government revenue and expenditure ratios for the set of EU-15 countries, at least at the 10% level of significance, both using conventional (asymptotic) critical values given in Pedroni (1999), and bootstrap panel cointegration proposed by Westerlund and Edgerton (2007). Moreover, this conclusion holds for the two sub-periods, 1970–1991 and 1992–2006 (broadly before and after the Maastricht Treaty), for most of the cointegration tests carried out.

Naturally, one has to stress that in this paper we assessed fiscal sustainability taking into account the stock of explicit government debt, and also via the analysis of cointegration relationships between the flows of government expenditures and revenues. Other aspects, outside the scope of analysis of the paper, and which are also relevant for the sustainability of public finances, are on the one hand the existence of implicit government liabilities, and on the other hand population ageing in combination with insufficiently funded public pension schemes that may endanger fiscal sustainability in the future.

Acknowledgments We are grateful to Ad van Riet, Jürgen von Hagen, participants at the 10th Banca d'Italia Public Finance Workshop (Perugia), at the 2008 EcoMod International Conference on Policy Modeling (Berlin), and to an anonymous referee for helpful comments and suggestions and to Simone Ruiz for assistance with the data. The opinions expressed herein are those of the authors and do not necessarily reflect those of the European Central Bank or the Eurosystem. Christophe Rault thanks the Fiscal Policies Division of the ECB for its hospitality. UECE is supported by FCT (Fundação para a Ciência e a Tecnologia, Portugal), financed by ERDF and Portuguese funds.

Appendix A

See Table 13.

Table 13 Summary of standard individual unit root test results

Set of results	First difference of stock of real government debt (2000 constant prices)		
1	Individual unit root tests	ADF, no unit root: *AT*, *FI*, *FR*, *UK*, *SW*	PP, no unit root: *AT*, *IR*, *LU*, *SW*
2	Individual unit root tests, with breaks	ZA, no unit root: *FI*, *UK*	

		General government revenue (% of GDP)	General government expenditure (% of GDP)
3	Individual unit root tests	ADF, no unit root: *AT*, *DE*, *LU*, *SW* PP, no unit root: *AT*, *FI*, *DE*, *LU*, *SW*, *UK*	ADF, no unit root: *DE*, *UK* PP, no unit root: *DE*, *LU*, *PT*, *UK*
4	Individual unit root tests, with breaks	ZA, no unit root: no countries	ZA, no unit root: *FI*, *FR*, *LU*, *PT*, *SP*

(a) *AT* Austria, *DE* Germany, *DK* Denmark, *FI* Finland, *FR* France, *GR* Greece, *IR* Ireland, *IT* Italy, *LU* Luxembourg, *NL* Netherlands, *PT* Portugal, *SP* Spain, *SW* Sweden, *UK* United Kingdom

(b) ADF refers to the Augmented Dickey–Fuller unit-root test; PP to the Phillips–Perron (PP) unit-root test (1988); ZA to the endogenous unit root tests by Zivot and Andrews (1992), allowing for one structural breaks

(c) For further details and explanation on the results of the conventional unit root tests we refer the reader to the extended working paper version in Afonso and Rault (2007)

Appendix B: panel unit root tests, additional results

See Tables 14 and 15.

Table 14 Panel data unit root tests for the government debt-to-GDP ratio (1970–2006)

Method	Statistic	*P*-value[a]	Cross-sections	Obs.
Null: unit root (assumes common unit root process)				
Levin, Lin and Chu t^a	0.54469	0.7070	15	525
Null: unit root (assumes an individual unit root process)				
Im, Pesaran and Shin W-stat.	1.68879	0.9544	15	525

Automatic selection of lags based on SIC. Newey–West bandwidth selection using a Bartlett kernel

[a] The tests assume asymptotic normality

Table 15 Panel data unit root tests for the primary balance-to-GDP ratio (1970–2006)

Method	Statistic	*P*-value[a]	Cross-sections	Obs.
Null: unit root (assumes a common unit root process)				
Levin, Lin and Chu t^a	−2.77057	0.0028	15	525
Null: unit root (assumes an individual unit root process)				
Im, Pesaran and Shin W-stat.	−4.34773	0.0000	15	525

Automatic selection of lags based on SIC. Newey–West bandwidth selection using a Bartlett kernel

[a] The tests assume asymptotic normality

References

Afonso, A. (2005). Fiscal sustainability: The unpleasant European case. *FinanzArchiv, 61*(1), 19–44.

Afonso, A. (2008). Ricardian fiscal regimes in the European Union. *Empirica, 35*(3), 313–334.

Afonso, A. & Rault, C. (2007). *What do we really know about fiscal sustainability in the EU? A panel data diagnostic* (Working Paper 820). Frankfurt a.M.: European Central Bank.

Ahmed, S., & Rogers, J. (1995). Government budget deficits and trade deficits. Are present value constraints satisfied in long-term data? *Journal of Monetary Economics, 36*(2), 351–374.

Banerjee, A. & Carrion-i-Silvestre, J. (2006). *Cointegration in panel data with breaks and cross-section dependence* (Working Paper 591). Frankfurt a.M.: European Central Bank.

Banerjee, A., Marcellino, M., & Osbat, C. (2004). Some cautions on the use of panel methods for integrated series of macro-economic data. *Econometrics Journal, 7*(2), 322–340.

Banerjee, A., Marcellino, M., & Osbat, C. (2005). Testing for PPP: Should we use panel methods? *Empirical Economics, 30*(1), 77–91.

Bergman, M. (2001). Testing government solvency and the no ponzi game condition. *Applied Economics Letters, 8*(1), 27–29.

Bohn, H. (1998). The behavior of U.S. public debt and deficits. *Quarterly Journal of Economics, 113*(3), 949–963.

Breitung, J., & Pesaran, M. (2005). Unit roots and cointegration in panels. In L. Matyas & P. Sevestre (Eds.), *The econometrics of panel data*. Boston: Klüver Academic Press.

Breuer, J., McNown, R., & Wallace, M. (2002). Series-specific unit root tests with panel data. *Oxford Bulletin of Economics and Statistics, 64*, 527–546.

Chang, Y., Park, J., & Song, K. (2006). Bootstrapping cointegrating regressions. *Journal of Econometrics, 133*(2), 703–739.

Choi, I. (2006). Combination unit root tests for cross-sectionally correlated panels. In D. Corbae, S. Durlauf, & B. Hansen (Eds.), *Econometric theory and practice: frontiers of analysis and applied research, essays in honor of Peter C. B. Phillips*. Cambridge: Cambridge University Press.

Claeys, P. (2007). *Sustainability of EU fiscal policies: A panel test* (Documents de Treball 2007/02). Institut de Recerca en Economia Aplicada, University of Barcelona.

Domar, E. (1944). The 'burden of debt' and the national income. *American Economic Review, 34*(4), 798–827.

Fève, P., & Hénin, P. (2000). Assessing effective sustainability of fiscal policy within the G-7. *Oxford Bulletin of Economic Research, 62*(2), 175–195.

Fisher, R. (1932). *Statistical methods for research workers*. London: Oliver and Boyd.

Gutierrez, L. (2006). Panel unit roots tests for cross-sectionally correlated panels: A Monte Carlo comparison. *Oxford Bulletin of Economics and Statistics, 68*(4), 519–540.

Hakkio, G., & Rush, M. (1991). Is the budget deficit 'too large?'. *Economic Inquiry, 29*(3), 429–445.

Hamilton, J., & Flavin, M. (1986). On the limitations of government borrowing: A framework for empirical testing. *American Economic Review, 76*(4), 808–816.

Haug, A. (1991). Cointegration and government borrowing constraints: Evidence for the United States. *Journal of Business & Economic Statistics, 9*(1), 97–101.

Im, K. & Lee, J. (2001). *Panel LM Unit Root Test with Level Shifts* (Discussion paper). Department of Economics, University of Central Florida.

Im, K., Pesaran, M., & Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics, 115*(1), 53–74.

Keynes, J. (1923). A tract on monetary reform. In *The collected writings of John Maynard Keynes*, vol. IV, London: Macmillan, (1971 edition).

Lee, J., & Strazicich, M. (2003). Minimum Lagrange multiplier unit root test with two structural breaks. *Review of Economics and Statistics, 85*(4), 1082–1089.

Levin, A., Lin, C.-F., & Chu, C.-S. (2002). Unit root tests in panel data: Asymptotic and finite sample properties. *Journal of Econometrics, 108*(1), 1–24.

Lyhagen, J. (2000). *Why not use standard panel unit root test for testing PPP* (Working Paper No. 413). Stockholm: Stockholm School of Economics.

MacDonald, R. (1992). Some tests of the government's intertemporal budget constraint using US data. *Applied Economics, 24*(12), 1287–1292.

McCallum, B. (1984). Are bond-financed deficits inflationary? A ricardian analysis. *Journal of Political Economy, 92*(1), 123–135.

McCoskey, S., & Kao, C. (1998). A residual-based test of the null of cointegration in panel data. *Econometric Reviews, 17*(1), 57–84.

Moon, H., & Perron, B. (2004). Testing for a unit root in panels with dynamic factors. *Journal of Econometrics, 122*(1), 8–126.

O'Connell, P. (1998). The overvaluation of purchasing power parity. *Journal of International Economics, 44*(1), 1–19.

Park, J. (2002). An invariance principle for sieve bootstrap in time series. *Econometric Theory, 18*(2), 469–490.

Payne, J. (1997). International evidence on the sustainability of budget deficits. *Applied Economics Letters, 12*(4), 775–779.

Pedroni, P. (1996). *Fully modified OLS for heterogeneous cointegrated panels and the case of purchasing power parity* (Working paper in economics No. 96-020). Bloomongton: Indiana University.

Pedroni, P. (1999). Critical values for cointegrating tests in heterogeneous panels with multiple regressors. *Oxford Bulletin of Economics and Statistics, 61*(1), 653–670.

Pedroni, P. (2000). Fully modified OLS for heterogeneous cointegrated panels. *Advances in Econometrics, 15*, 93–130.

Pedroni, P. (2004). Panel cointegration; asymptotic and finite sample properties of pooled time series tests with an application to the purchasing power parity hypothesis. *Econometric Theory, 20*(3), 597–625.

Pedroni, P., & Urbain, J.-P. (2001). *Cross member cointegration in non-stationary panels*. Mimeo: Universtiteit Maastricht.

Pesaran, M. (2004). *General diagnostic tests for cross section dependence in panels* (Cambridge Working Papers in Economics, No. 435), University of Cambridge.

Phillips, P. (2001). *Bootstrapping spurious regressions* (Cowles Foundation Discussion Paper No. 1330). Yale.

Phillips, P., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika, 75*(2), 335–346.

Phillips, P., & Sul, D. (2003). Dynamic panel estimation and homogeneity testing under cross-section dependence. *Econometrics Journal, 6*(1), 217–259.

Prohl, S. & Schneider, F. (2006). *Sustainability of public debt and budget deficit: Panel cointegration analysis for the European union member countries* (Working Paper No. 0610). Department of Economics, Johannes Kepler University Linz.

Quintos, C. (1995). Sustainability of the deficit process with structural shifts. *Journal of Business & Economic Statistics, 13*(4), 409–417.

Schmidt, P., & Phillips, P. (1992). LM tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics, 54*, 257–287.

Tanner, E., & Liu, P. (1994). Is the budget deficit 'too large'? Some further evidence. *Economic Inquiry, 32*, 511–518.

Taylor, M., & Sarno, L. (1998). The behavior of real exchange rates during the post-bretton woods period. *Journal of International Economics, 46*(2), 281–312.

Taylor, A., & Taylor, M. (2004). The purchasing power parity debate. *Journal of Economic Perspectives, 18*(4), 135–158.

Trehan, B., & Walsh, C. (1991). Testing intertemporal budget constraints: Theory and applications to U.S. federal budget and current account deficits. *Journal of Money Credit and Banking, 23*(2), 206–223.

Uctum, M., & Wickens, M. (2000). Debt and deficit ceilings, and sustainability of fiscal policies: An intertemporal analysis. *Oxford Bulletin of Economic Research, 62*(2), 197–222.

Westerlund, J. & Edgerton, D. (2007). A panel bootstrap cointegration test. *Economics Letters*, *97*(3), 185–190.

Wilcox, D. (1989). The sustainability of government deficits: Implications of the present-value borrowing constraint. *Journal of Money Credit and Banking, 21*(3), 291–306.

Zivot, E., & Andrews, D. (1992). Further evidence of the great crash, the oil-price shock and the unit-root hypothesis. *Journal of Business and Economic Statistics, 10*(3), 251–270.